

**THE EFFECT OF GENDER STEREOTYPES IN LANGUAGE
ON ATTITUDES TOWARD SPEAKERS**

by

Christy L. Dennison

B.A., English, Chatham College, 1999

Submitted to the Graduate Faculty of
Arts and Sciences in partial fulfillment
of the requirements for the degree of
Master of Arts in Applied Linguistics

University of Pittsburgh

2006

UNIVERSITY OF PITTSBURGH

ARTS AND SCIENCES

This thesis was presented

by

Christy L. Dennison

It was defended on

December 15, 2005

and approved by

Shelome Gooden, Ph.D., Assistant Professor, Linguistics

Amanda Godley, Ph.D., Assistant Professor, Instruction & Learning

Thesis Advisor: Scott F. Kiesling, Ph.D., Assistant Professor, Linguistics

THE EFFECT OF GENDER STEREOTYPES IN LANGUAGE ON ATTITUDES TOWARD SPEAKERS

Christy L. Dennison, M.A.

University of Pittsburgh, 2006

This study uses a matched guise technique to elicit evaluations of men and women from participants based solely on what they hear. Four speakers (two men and two men) created two recordings, one in which they incorporated “women’s language” into their speech and the other using “standard” language. One hundred university students listened to each recording and evaluated the speaker in terms of twelve personality traits. Results showed a significant difference in how male and female speakers were perceived, regardless of the language style they employed. Women’s language and “standard” language were also perceived differently regardless of speaker gender.

TABLE OF CONTENTS

1.0	INTRODUCTION.....	1
2.0	REVIEW OF RELEVANT LITERATURE.....	3
3.0	METHODOLOGY.....	13
3.1	PARTICIPANTS	13
3.2	MATERIALS	14
3.3	PROCEDURE	17
4.0	RESULTS AND DISCUSSION	19
4.1	FACTOR ANALYSIS	19
4.2	PERCEPTION OF SPEAKERS	21
4.3	THE MATCHED GUISE.....	22
4.4	DIMENSIONS SCORES.....	30
4.4.1	Masculinity	30
4.4.2	Femininity	32
4.4.3	Nurturance.....	34
4.5	EFFECT OF PARTICIPANT GENDER	41
5.0	CONCLUSION.....	45
5.1	DIRECTIONS FOR FUTURE RESEARCH	46
	APPENDIX A	49
	APPENDIX B	52
	BIBLIOGRAPHY	53

LIST OF TABLES

Table 1: Matched Guise Recording Content by Speaker.....	15
Table 2: Factor Matrix (Omitting "Analytical" and "Individualistic")	21
Table 3: Comparing Perceptions of Same Speaker When Using Different Language Styles	23
Table 4: Comparing Perceptions of Different Speakers (A and C) Using Same Language Style.....	27
Table 5: Comparing Perceptions of Different Speakers (B and D) Using Same Language Style.....	28
Table 6: Summary of Two-way (speaker gender X speaker language style) ANOVA on Masculinity Scores.....	30
Table 7: Means and Standard Deviations of Masculinity Scores by Speaker Gender and Language Style.....	31
Table 8: Summary of Two-Way (Speaker Gender X Speaker Language Style) ANOVA on Femininity Scores	32
Table 9: Means and Standard Deviations of Femininity Scores by Speaker Gender and Language Style	33
Table 10: Summary of Two-way (speaker gender X speaker language style) ANOVA on Nurturance Scores.....	34
Table 11: Means and Standard Deviations of Nurturance Scores by Speaker Gender and Language Style.....	35

LIST OF FIGURES

Figure 1: Comparison of Masculinity Scores for Each Speaker by Language Style.....	24
Figure 2: Comparison of Femininity Scores for Each Speaker by Language Style	25
Figure 3: Comparison of Nurturance Scores for Speakers Based on Language Style.....	26
Figure 4: Comparison of Masculinity Scores by Gender and Language Style.....	31
Figure 5: Comparison of Femininity Scores by Gender and Language Style	33
Figure 6: Comparison of Nurturance Scores for Speaker Gender by Language Style	35
Figure 7: Comparison of Masculinity Scores for Men Using Women's Language and Women Using "Standard" Language Based on Participant Gender	42

1.0 INTRODUCTION

Robin Lakoff's *Language and Woman's Place (LWP)*, published in 1975, was one of the first publications of its time to address the relationship between language and gender. As a result, *LWP* served as the impetus for both linguists and feminists alike to look more closely at gender variation in language. Many studies expanded on Lakoff's argument that language embodies gender inequity. As stated in *LWP*, there is a direct correlation between the inequity in language and the inequity in men and women's social statuses. "Women's language", a term coined by Lakoff, became a commonly used identifier among language and gender researchers. According to Lakoff, women's language describes how women use language and how language is used to talk about women, both which position women as powerless. This position has been adopted by a number of sociolinguistics and feminists. In an effort to support Lakoff's rationale with empirical evidence (Lakoff's argument was based entirely on personal observation), many studies focused on identifying and quantifying the linguistic resources that men and women utilize when they speak. The overall goal of such studies has often been the classification of certain linguistic features as occurring more often in women or in men. However, other researchers have found that gender variation in language is not as clearly and easily defined as much of the earlier research may claim. The influence of context (local and global), social factors other than gender (ethnicity, age, socioeconomic status), and issues of power have also been found to play a role in how men and women use language.

Although it takes a different approach to the sociolinguistic exploration of gender inequity in language, this thesis incorporates much of the work that linguists like Lakoff have contributed to the field. Instead of focusing solely on speakers and the linguistic resources they employ, this study also investigates how speakers are evaluated by listeners. The argument for

such an approach regards it as insufficient to invest all of our attention into how gender stereotypes are perpetuated by the speaker alone (for example, because a woman uses linguistic feature X, she is, as Lakoff would claim, weak or ineffectual). It is also necessary to consider a listener's attitude toward the man or woman speaking, as the interpretation of speech can have significant implications for the nature of social relationships. For example, Delph-Janiurek (1999) found that the manner in which college professors spoke had a profound effect on how students evaluated them, specifically regarding gender. Voice qualities such as a lack of enthusiasm and emotion were considered by students to be "masculine" (in men and women), whereas instructors (both men and women) with more "emotional" vocal performances that involved breathiness or a higher pitch were viewed as "feminine".

Along these same lines, this thesis addresses how men and women are evaluated (in terms of personality traits) by listeners. Taking into account the gender-specific functions that have been attached to language (as first posited by Lakoff), this thesis asks the following questions: first, is there evidence that these gender-specific functions play a role in how a listener perceives a speaker? In other words, does a listener's evaluation of a speaker demonstrate an awareness of stereotypically gendered language characteristics, such as women's language? And second, if we elaborate on the first question posed, would a woman using women's language be subject to the same evaluation as a man using women's language?

2.0 REVIEW OF RELEVANT LITERATURE

The research questions posed in this thesis address Lakoff's claim in *LWP* that certain linguistic features have very specific gendered connotations. In her discussion of women's language, Lakoff describes features such as tag questions ("this election mess is terrible, *isn't it?*"), rising intonation on declarative statements (responding to a request for the time with "*Six o'clock?*") and hedges ("That's *kinda* sad", or "I'm *sort of* angry with you"), all of which function to mitigate a woman's position. For example, according to Lakoff's argument, a tag question usually indicates a speaker's uncertainty or lack of commitment to what is being said. It would follow then, based on Lakoff's theory, that women use more tags than men since they are "weaker" and, thus, less likely to make an unmitigated statement. However, this did not always turn out to be the case empirically.

Because Lakoff's claims were based solely on her own observations and introspection, many sought to support her statements with empirical evidence. Fishman (1980), a well known study about women's interactional "work", focused on how the verbal interaction between intimate heterosexual couples created and maintained the hierarchical social relationship between men and women. Positioned within Lakoff's framework that identifies certain linguistic resources as functioning solely to weaken the speaker's position, Fishman's analysis suggests that women work harder than men in interaction, and that interaction between men and women is most often on men's terms. The data consisted of over fifty hours of interaction between couples

in their homes, and revealed that women ask more questions and use more devices described as insuring rights to speak (saying “D’ya know what?” at the beginning of a conversation) and establishing interest (“This is interesting” as an introduction to a topic) than men. Furthermore, according to Fishman, men and women used minimal responses (using words such as “*uh-huh*” and “*right*” while someone is talking - also known as backchanneling) quite differently – for women, it is “support work” used to indicate that they are attending to their male partner’s speech. But for men, Fishman states, the use of this device shows a lack of interest in what the woman is saying. Additionally, Fishman found that men made more statements that received a response (such as a lengthy conversation) than women did. Many, if not most, of women’s statements did not get a response from men which, for Fishman, implies that men only engage in conversation on their own terms.

While it is ideal to investigate language use within real interaction as Fishman did, as opposed to a context-free vacuum, the analysis of Fishman’s data is problematic. Details about the context in which the interactions occurred are completely disregarded. In her analyses, Fishman mapped salient gender stereotypes (for example men are dominant, while women are childlike) onto linguistic forms and provided no other possibilities for interpretation of these interactions. It was revealed that women used devices to insure their right to speak twice as often as men and to establish interest in a subject thirty-four times compared to men’s three times; however, men did employ these devices (albeit less often), so it cannot be said that *only* women used them. Unfortunately, the analysis gives no explanation as to the function these devices had for men within the interactions, regardless of how infrequently they occurred. Instead, Fishman focused solely on frequency and argued that since women used these devices more often than men, they must function to keep women “in their place” (that is, in a powerless

position). Information about the backgrounds of the participants – though provided in the description of the study – was not factored into the analysis of the data. This is not to say that the frequency is insignificant; in fact, quite the contrary is true. It is essential to language and gender research to explore which linguistic features, if any, are used more often by men or women. Nevertheless, it is also crucial to examine closely the social contexts in which interactions occur, including the backgrounds of the participants and the nature of the relationship. Otherwise, the analysis is missing a large part of the story behind why a person uses the language he or she uses.

Much of the early analyses of discourse between men and women interpreted data through a lens that already perceived women as powerless and men in control. As a result, whatever women do in interaction – whether it is more or less often than men – is explained in terms of how it perpetuates women's subordinate position. This interpretation assumes homogeneity of the gender category and omits the significance of other factors that play into one's social identity (socioeconomic status, race, age, etc.). As previously noted, the main strength of research like Fishman (1980) is the use of real interaction for the collection of data. It takes Lakoff's claims one step further by applying them to the real world and measuring them quantitatively; however, we also see how such analyses can limit the possible interpretations by not considering social context and participant background, as well as the multifunctionality of linguistic features.

Unfortunately, there is little point in collecting data in a "real" context if that context is going to be disregarded in the interpretation. To address this, several studies introduced alternative interpretations of the features Lakoff labeled "women's language" (Cameron, *et al.*

(1988); O'Barr and Atkins (1980); Zimmerman and West (1975; 1983))¹. In an article entitled "Lakoff in context: social and linguistic functions of tag questions", Cameron *et al.* (1988) examine numerous functions that a tag question can have, depending on its form and context. Citing the now obvious weakness of *LWP*'s lack of empirical evidence, Cameron *et al.* conducted a study based on the analysis of Holmes (1984) which identified two main functions of tag questions: modal and affective.² The objective was to prove that Lakoff's definition of the mitigating function of tags is too restrictive and that tags do not necessarily signify tentativeness or weakness. In their attempt to label the tags found in their data, Cameron *et al.* discovered the problematic nature of such labels. In actuality, several of the tag examples in their data appeared to serve many functions simultaneously. Thus, the study suggests that the link between linguistic function and form is not invariant, as both Lakoff and Fishman have implied. Additionally they stress the importance of considering factors other than gender when analyzing linguistic patterns:

These include the role taken by participants in interaction, the objectives of the interaction, participants' relative status on a number of dimensions, and so on...Gender is cross-cut with other social divisions and their relative importance is affected by the specifics of the situation (p.47).

Also interesting in this study is the implication that the use of a tag as a conversation facilitator – that is, as a way to initiate a response from someone – is actually a marker of conversational control, rather than a device that a subordinate speaker uses to keep the conversation going. In

¹ The linguistic features that Lakoff originally identified as components of women's language remained the focus of much research and analysis well after the publication of *LWP*, despite some arguments that these features were "arbitrary" (Cameron et al., 1988).

² For a more in-depth explanation of the different functions of tag questions, see Holmes (1984).

fact, Cameron *et al.* suggest that use of such a linguistic device could be a way of coping with oppressive conditions or even resisting them.

What studies like this reveal is that it is problematic to only consider subordination and weakness when identifying women's language. It is more probable that there are other elements playing a role in speech patterns. O'Barr and Atkins (1980), in their study of the speech of courtroom witnesses, propose that women's language is actually powerless language, and that many of the linguistic features identified by Lakoff as components of women's language are in fact used by people (men and women) who are in a socially powerless role within a specific context (witness in a courtroom, for example). Based on more than 150 hours of recorded testimony from various witnesses, they discovered that not all women use Lakoff's features and that some men do, and that it often depended on one's socioeconomic status, experience, and/or occupation. Thus, so-called "women's language" is a reflection of social position, and "using this type of language...tends to feedback into the social situation" (p. 110). So, once again, we see a divergence from the original hypothesis that Lakoff presents in *LWP*, and the complexity of the interaction between language, social context, and participants becomes more evident.

The value of Lakoff's argument is certainly not being disputed in the studies reviewed here. Rather, their results exemplify how essential the publication of *LWP* was in serving as an impetus for the study of language and gender. More questions have been asked about the relationship between language and gender, and as we attempt to answer them, it becomes apparent that the answers are not simple. With each "phase" of research, new elements are being considered in terms of their influence on language choices and patterns. This leads to more sophisticated and intricate interpretations. In fact, several studies have taken the work of Lakoff, Fishman, etc. even further by adding another dimension to analysis – perception (Erickson *et al.*

(1978); Batstone & Tuomi (1981); Strand (1999); Delph-Janiurek (1999; 2000)). *Perception* in all of these cases except for Strand (1999), and for purposes of this thesis, refers to an “outside” listener’s interpretations of a speaker’s language and/or personality based (primarily) on the speaker’s language style. In Strand (1999), however, the term *perception* refers to the cognitive processing of acoustic cues from speech input and does not involve the evaluation of the speaker’s personality.³

In one of the earliest studies using perception to explore the relationship between language and gender, Erickson *et al.* (1978) found that, in a courtroom context, speech incorporating women’s language (or “powerless language” as they refer to it) was evaluated negatively regardless of the speaker’s gender. Batstone & Tuomi (1981) discovered that men and women identified the same characteristics in women’s speech, but rated these characteristics differently in terms of salience. In a more recent study, Strand (1999) explored how certain “triggered” judgments about a speaker play a role in the perception of language.

Conducting a phonetic experiment, Strand considered the gender-related variability in the production of the fricatives [s] and [ʃ] for the phoneme /s/. Though there is little physiological evidence for the difference between men and women’s production of this sound (women have been observed as producing the voiceless alveolar fricative in different manners – women’s production is closer to [s], where as men’s production is more like [ʃ]), Strand acknowledges that something is factoring into the variation between the two. She argues that children are socialized to produce the form “appropriate” for their gender. Furthermore, listeners, based on such socialization, have very specific expectations of speakers and organize the input as quickly

³ In an effort to avoid confusion, it should be stated that in this thesis the term *perception* will often be used interchangeably with words like *attitude* and *evaluation* to refer to people’s assessments of speakers.

as possible, based on stereotypes. What implications does this have for gender? Geis (1993, as cited in Strand 1999) states that stereotypes

enhance perceptions, interpretations and memories that are consistent with stereotypical attributes and obscure, diffuse, or cause us to disregard or forget information that is inconsistent with them...Thus, even when women and men behave alike, we see them as different (p. 95).

As it is argued in a number of studies, the term “women’s language” becomes problematic when evidence shows not just that men use it in certain contexts, but that certain linguistic features have various functions. So, Strand asks if gender stereotypes play a role in how features of speech are perceived. And, in fact, her results revealed that the same fricative (/s/ and /ʃ/) was perceived differently, depending on whether it was matched with the image of a man or woman’s face. She suggested that listeners incorporated gender stereotypes into their perceptions; that is, they made decisions about what they heard based on what they think the speech of men and women “should” sound like. Clopper *et al.* (2005) also found that gender-specific information played a role in how listeners categorized the dialects of speakers. Listeners used their own expectations regarding gender to identify the dialects of the speakers they were hearing. It was found that linguistic variation based on gender (for example, perceived differences between men’s and women’s vowel production) assisted listeners in classifying the dialects.

The issue at the core of these studies - namely, the role of stereotypes in the perception of language and language users – is central to this thesis as well. This study takes into consideration the claims of studies that emphasize the link between linguistic form and gender as well as the perceptions and attitudes of listeners. The objective is a better understanding of the awareness (or lack thereof) a listener has of these gendered functions of certain linguistic forms. According to Lakoff, the features that make up women’s language are predominantly, if not

solely, used by women; therefore, it would logically follow that men who use tag questions would be considered to be feminine, or, at the very least, possess feminine traits. Yet this concept has not been extensively explored in the research. O'Barr and Atkins (1980) addressed the fact that men within a courtroom context (as witnesses) use features identified as women's language; however, the male speakers in this study were not perceived in terms of their femininity and masculinity. Rather, O'Barr and Atkins suggested using the term "powerless language" to refer to this speech style, which removes the focus from the issue of gender and emphasizes context.

More recent studies like Strand (1999) and Clopper (2005) address gender with a compelling argument regarding the important role that stereotypes play in the perception of language. These studies have provided an interesting direction for investigating how women's language is perceived, which could result in enhanced understanding of gender variation in language. The study conducted for this thesis uses listener perception to explore the issues raised by Lakoff and her contemporaries. First, it identifies femininity and masculinity in terms of culturally defined personality traits (according to the Bem Sex Role Inventory, which is discussed in detail in the methodology section). It then examines the extent to which these stereotypical traits are perceived in women and men using both women's language and a more standard style of speech.⁴ For example, because listeners are hearing a woman using women's language, they may give her high rankings for the presence of *shyness*, *warmth*, and *sensitivity*; yet a man using the same language could be given low scores for these traits. This would suggest that when hearing a woman speak, people associate certain qualities with her based on

⁴ *Standard*, for purposes of this study, refers to speech lacking the specific characteristics that Lakoff defined as features of women's language. It is acknowledged that identifying a single, standard form of language is near to impossible. Thus, it may be more accurate to state that this study utilizes a form of language that – compared to women's language – is *more similar to a standard style of language*.

gender stereotypes. Just as Strand (1999) revealed, there is evidence that if people are socialized to have certain subconscious expectations of how a man and woman should be (or should speak), they are likely to attribute to the speaker those traits that fit their expectations. If men and women are given similar evaluations, we could infer that these stereotypes are not influencing perceptions and that there is something else/more playing a role, such as an awareness of the multiple functions of a certain linguistic feature. This would also suggest that the function of women's language is not invariant, as Lakoff and many of her contemporaries have suggested.

The main hypothesis for this thesis is that speaker gender will have a more significant effect than the style of language the speaker uses on how the speakers are evaluated. This falls in line with the findings of Strand (1999), which are explained in terms of the "very specific expectations of speakers" that listeners have based on gender. As Strand puts it, listeners use stereotypes to "organize the input" as quickly as possible. In this study, it is expected that the listener-judges, upon hearing a man or woman speaker, would categorize what they hear based on how they have been socialized with regard to gender and what they know about how men and women "should" speak and behave.

The relationship (if any) between the gender of the listener-judges and the nature of their responses for each speaker and language style will also be explored to determine whether men and women rate the speakers differently. The speaker/language style combinations to be considered include: men using women's language, men using standard language, women using women's language, women using standard language, as well as men and women overall. Here, there are two possible outcomes: there will be no significant difference between how men and women rate the speakers, or there will be a significant difference between men's and women's evaluations of the speakers, although exactly how these evaluations would differ remains to be

seen. The hypothesis predicts that there will be no significant difference between how men and women evaluate the speakers. This prediction is based on the claim (which is supported by Strand (1999)) that people are socialized similarly when it comes to gender expectations. That is, both men and women have basically the same idea about how men and women “should” speak and behave. Therefore, all the listener-judges, regardless of their own gender, will evaluate men and women speakers similarly.

3.0 METHODOLOGY

3.1 PARTICIPANTS

One hundred undergraduate and graduate students at the University of Pittsburgh participated in this study in exchange for “extra” academic credit for the course in which they were recruited. Participants were recruited from five different linguistics classes (two sections of Introduction to Linguistics, two sections of Phonetics & Phonemics, and one section of Language, Gender & Society). If students were interested in participating in this study for extra credit, they were asked in class to sign up for one of three scheduled “listening sessions” (details about this later in this section) using their email address only for reasons of confidentiality. Four non-native English speakers (three women and one man) participated in the study; however, because the goal of the study was to examine stereotypical attitudes regarding gender, which are typically culturally prescribed, the non-native English speakers were omitted from the analyses. Also not included in analysis were eight participants whose responses, as a result of an undetermined computer or software malfunction, were lost. Therefore, the total number of participants used in the analyses was 88 (58 women and 30 men, all native speakers). Ages of participants ranged from 18 to 46 years old ($M = 21.75$).

3.2 MATERIALS

To investigate the research questions posed, a *matched-guise* technique was used, which measures language attitudes. An excellent example of the use of this technique is Lambert (1967), which utilized a matched-guise to compare attitudes and biases toward French-speaking and English-speaking Canadians. In that study, listeners were presented with recordings by bilingual speakers who each recorded two passages – one in French and one in English. Listeners were not aware that they were listening to the same person using two different languages. They were then asked to fill out a questionnaire about their attitude toward the speaker in each recording. Lambert (1967) discovered that, despite the fact that listeners were hearing the same person, the recordings created in French elicited many more stereotypical and negative perceptions.

Applying this concept to my research, eight recordings were created with the assistance of four graduate students in the Theatre Arts department at the University of Pittsburgh. Given their experience in adjusting and manipulating their speech, it was believed that using theatre students would lend to the authenticity of the recordings. Two men and two women were employed to create two recordings each – one incorporating features of women’s language (as Lakoff described it) and the other devoid of such features. Features of women’s language included were: high rising intonation in declaratives, hedges (*kinda, sorta, a little*), intensifiers (*really, so, very*), and qualifying remarks⁵ (beginning a statement with ‘*I think maybe*’ or ‘*I don’t know about this, but...*’). Each recording was in the form of an answering machine message by which the speaker inquired about an apartment for rent. The speaker was given a page from the classifieds section of the newspaper and told to select a rental ad. The only requirements given

⁵ Qualifying remarks are examined in Kramer (1974b); Siegler & Siegler (1976), as referenced by Kramarae (1982).

to the speaker were the types of features to include when using women's language. Each of these recordings incorporated two types of features (for example, high rising intonation and intensifiers or hedges and qualifying remarks). They were also coached briefly regarding the differences between women's language and a more standard language so they could adjust their speech accordingly in both recordings. For example, the speakers were told to completely eliminate all high rising intonation in declarative statements from their standard language recording, so as to sound more certain. They were not given a script to ensure a more authentic and less practiced sounding recording, although most of them did take notes about what they planned to say. However, linguistic features such as fillers and hesitations were not discouraged. Average length of the recordings was 20.25 seconds. Transcriptions of each recording can be found in Appendix A. The recordings were created using the computer program Audacity version 1.2.3 (audio editor and recorder). Table 1 summarizes the content (in terms of linguistic features) of each of the eight recordings, including the number of tokens of each feature (as indicated by the number in parentheses). The content of standard language recordings is not explained because these recordings were simply devoid of any features typically associated with women's language. (For a better understanding of the content for standard recordings, the transcriptions in Appendix A can be reviewed.)

Table 1: Matched Guise Recording Content by Speaker

	<u>Recording #1</u>	<u>Recording #2</u>
Speaker A (male):	High rising intonation (4), intensifiers (3)	Standard language
Speaker B (female):	Qualifying remarks (2), hedges (3)	Standard language
Speaker C (male):	Qualifying remarks (2), hedges (1)	Standard language
Speaker D (female):	High rising intonation (9), intensifiers (4)	Standard language

After listening to each recording, participants were asked to complete a Likert-scaled survey in which they evaluated that particular speaker based on twelve adjectives. Each adjective

represented a personality trait, and the extent to which the trait was exhibited by the speaker was ranked by listeners on a scale of one to five, one being no presence of the trait for the speaker and five indicating a very strong presence of the trait. The adjectives were adapted from the Bem Sex Role Inventory (BSRI), developed in 1974 to measure masculinity and femininity as they are defined culturally. The judges used to screen the adjectives for the BSRI were asked to rate each trait in terms of how desirable it was for either a man or a woman in American society. According to Bem (1974), the traits considered to measure masculinity have in common a “cognitive focus on getting the job done” (examples include *self-reliant*, *assertive*, and *willing to take risks*), while the traits that purport to measure femininity involve an “affective concern for the welfare of others” (examples include *yielding*, *cheerful*, and *compassionate*). The BSRI has been the subject of a number of studies questioning its validity in measuring gender traits, and while defining gender in these terms is inarguably problematic, the adjectives were suitable for the purposes of this survey not in spite of their stereotypic connotations, but because of them. In fact, the goal of the BSRI was not to fit test takers into neat gender categories, but rather to identify cultural definitions of gender appropriateness and to determine to what extent an individual works to maintain a culturally appropriate gender identity. Therefore, because the adjectives used in the inventory tap into cultural stereotypes, they are quite useful for this thesis study which also addresses collective definitions of gender roles and expectations.

The survey was transformed into a computer interface so that participants could complete it on a personal computer in the Robert Henderson Language Media Center at the University of Pittsburgh. The survey was made up of ten screens. The first page elicited participant background information, the following eight screens were identical (except for the sound file) and contained the twelve adjectives and rankings used in the evaluation of each recording, and

the final page concluded the survey. Participant responses were then stored in a text file and saved on the university server, which is password protected. Once all participants completed the survey, the data were transferred to a compact disc by Media Center staff and given to the researcher to conduct analyses. Data were then removed entirely from the server.

3.3 PROCEDURE

As previously stated, students were recruited in five different linguistics classes. Professors provided students with a one-page handout explaining the purpose of the study and any benefits or risks resulting from their participation. Although they were not made aware of the focus on gender in the experiment to avoid any biases, the handout did explain that the purpose of the study was “to explore the relationship between how a person speaks and how that person is perceived by others”. The scheduled listening sessions were also provided so that interested students could choose when they want to participate.

When participants arrived to one the three scheduled listening sessions at the media center, they were asked to sign in using only their email and the course for which they wish to receive extra credit. Each participant was given a form with a unique identification number based on the order in which they arrived for the session. An example of this form can be found in Appendix B. They were directed to a computer in the lab (with a set of headphones) and told to read the form before beginning the task. The form also included a brief, written explanation of the task that encouraged participants to compare their task to evaluating an unfamiliar person they hear over the telephone or on an answering machine (which is not unusual to do). They were asked to answer using their first impressions of the speaker, and not to think too long and

hard about their responses. Each page of the computerized survey included one sound file and twelve adjectives to be considered in the evaluation of the speaker's personality. After giving brief details regarding their background (age, gender, native language, class status, major or graduate program), participants moved on to the survey. They clicked the mouse on the "play" button and listened to the recording. Once it ended, they were able to indicate how often the speaker displayed a given trait (never, rarely, sometimes, frequently, always). Participants could listen to the recording as often as they wanted, but while it played they could not click on any other buttons on the page. They were able to complete the survey at their own pace. When they finished evaluating the eighth and final recording, they were shown a screen thanking them for their participation and were then free to leave. Most participants were able to complete the entire survey in approximately 20 minutes.

4.0 RESULTS AND DISCUSSION

4.1 FACTOR ANALYSIS

To identify the dimensions underlying participants' ratings of the recordings, a principal axis factor analysis was conducted on the twelve adjectives used in the questionnaire. The goal of the factor analysis was also to determine if the adjectives actually coincided with the “masculine” and “feminine” labels attached to them in the BSRI. For the analysis, the only recordings selected were those of men using standard language (recording #1 and #5) because of the potential for variation of the factor structure across speaker gender and language style. Therefore, men using standard language were considered to be, for purposes of the factor analysis, the standard in terms of gender and language style. Using men speaking standard language as the comparison group makes sense considering that much of the research guiding this study – Lakoff included – identify men as the standard and women as the “other” whose language deviates from that standard.

Initially, all twelve adjectives were submitted to the factor analysis procedure. However, based on preliminary results, the adjectives *analytical* and *individualistic* were omitted because of low communalities (meaning these terms did not cluster with any of the other adjectives or with each other). When the remaining ten adjectives were submitted, two factors were extracted. The first factor appeared to capture the masculine/feminine dimension: the adjectives *independent*, *dominant*, *masculine*, and *assertive* had high (greater than .4) positive loadings,

whereas the adjectives *shy*, *childlike*, and *feminine* had high negative loadings. For the second factor, the adjectives *sensitive*, *warm*, and *understanding* all had high positive loadings. These results were not surprising as a number of factor analyses conducted with the adjectives in the BSRI have revealed that Bem's "unidimensional" categories (masculine and feminine) were not adequate (Brems and Johnson, 1990). Many of these analyses identified a multidimensional factor structure, and used labels such as Emotional Expressiveness (Moreland et al (1978)) or Interpersonal Sensitivity (Pedhazur and Tetenbaum (1979)) to describe the dimensions revealed by their analyses.

Because the adjectives that did not fit in the masculinity-femininity dimension all seem to represent traits that focus on nurturing qualities, the second dimension was labeled "nurturance". Thus, two dimensions were captured by the ten adjectives submitted to the factor analysis: "masculinity-femininity" and "nurturance". Scores were then created for the participants' evaluations of speakers based on three different dimensions: masculinity, femininity, and nurturance. Masculinity scores were determined by averaging ratings for the adjectives *independent*, *dominant*, *masculine*, and *assertive*; femininity scores were determined by averaging ratings for the adjectives *shy*, *childlike*, and *feminine*; and, finally, nurturance scores were determined by averaging ratings for the adjectives *warm*, *sensitive*, and *understanding*. Table 2 shows the results of the factor analysis. Adjectives with a loading lower than -.500 made up the masculinity-femininity factor and adjectives with a positive loading higher than .500 were considered to be part of the nurturance factor.

Table 2: Factor Matrix (Omitting "Analytical" and "Individualistic")

Rotated Factor Matrix^a

	Factor	
	1	2
Shy	-.624	-.072
Independent	.567	-.008
Dominant	.757	-.076
Sensitive	-.453	.598
Childlike	-.517	.248
Masculine	.699	-.128
Assertive	.590	-.179
Warm	-.063	.805
Feminine	-.607	.283
Understanding	-.007	.651

Extraction Method: Principal Axis Factoring.

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 3 iterations.

4.2 PERCEPTION OF SPEAKERS

As outlined above, the three dimensions were used to create new scores for each participant and each of the eight recordings (thus resulting in eight scores for each participant). The scores for each pair of recordings containing speakers of the same gender using the same language style (for example, recording #1 and #5, both of which demonstrated men using standard language) were then averaged, creating four summary scores on each of the three dimensions. The summary scores represent averages for male speaker/ standard language, male speaker/women's language, female speaker/ standard language, and female speaker/women's language.

The issue of analyzing ordinal data using parametric statistics (summary scores) was addressed while considering how to best analyze the results of this survey. It was determined

that for the purposes of this study, using parametric statistics is the most suitable method. Forming scores for the specific dimensions (masculinity, femininity, and nurturance) by averaging each participant's response better conforms to the properties expected when using parametric statistics than simply rating each adjective. For example, although it cannot be said that the difference between *rarely* and *sometimes* on the Likert scale is the same as the difference between *sometimes* and *frequently*, it is important to consider that a rating for each adjective would not be as useful for the intentions of this study. Rating each adjective would only allow a limited number of values (1 to 5), but using the sum has a larger range (5 to 60). The decision to use parametric statistics in this analysis is supported by Glass & Hopkins (1996), who state that "the particular scale of measurement is influenced by the interpretation to be drawn from the data" (10).

4.3 THE MATCHED GUISE

Before describing the main results of this study, it is important to outline first the major details of the matched guise technique. Various t-tests were conducted to compare scores for the same speaker using each of the two language styles (women's language, standard language). Scores were generated for each of three dimensions extracted from the factor analysis (masculinity, femininity, and nurturance). Overall, the matched guise used in this study was successful. Significant differences in perception occurred when a speaker used women's language and when he/she used standard language. For example, Speaker A (male) using standard language was evaluated as more masculine ($M = 3.76$), less feminine ($M = 2.00$), and less nurturing ($M = 2.42$) than Speaker A using women's language ($M = 2.95, 2.80, 3.73$ for each dimension, respectively).

All four speakers were considered more masculine when using standard language, more feminine when using women's language, and more nurturing when using women's language. This is shown in Table 3, which summarizes the results of all of the t-tests. Figures 1, 2, and 3 demonstrate the comparisons for each dimension graphically.

Table 3: Comparing Perceptions of Same Speaker When Using Different Language Styles

<u>Speaker</u>	<u>Dimension</u>	<u>"Standard"</u>		<u>Women's</u>		<u>T</u>
		<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	
A (Male)	Masc.	3.76	0.66	2.95	0.59	8.80*
	Fem.	2.00	0.71	2.80	0.65	-8.95*
	Nurt.	2.42	0.59	3.73	0.49	-16.74*
B (Female)	Masc.	3.32	0.63	2.37	0.70	9.73*
	Fem.	2.76	0.55	3.30	0.66	-6.23*
	Nurt.	3.01	0.66	3.15	0.60	-1.70
C (Male)	Masc.	3.61	0.63	2.72	0.48	13.06*
	Fem.	2.07	0.60	2.74	0.64	-9.10*
	Nurt.	3.15	0.53	3.47	0.53	-4.52*
D (Female)	Masc.	3.14	0.57	2.29	0.51	-10.23
	Fem.	2.86	0.57	3.53	0.52	10.09
	Nurt.	3.39	0.62	3.96	0.65	6.93

* $p < .001$

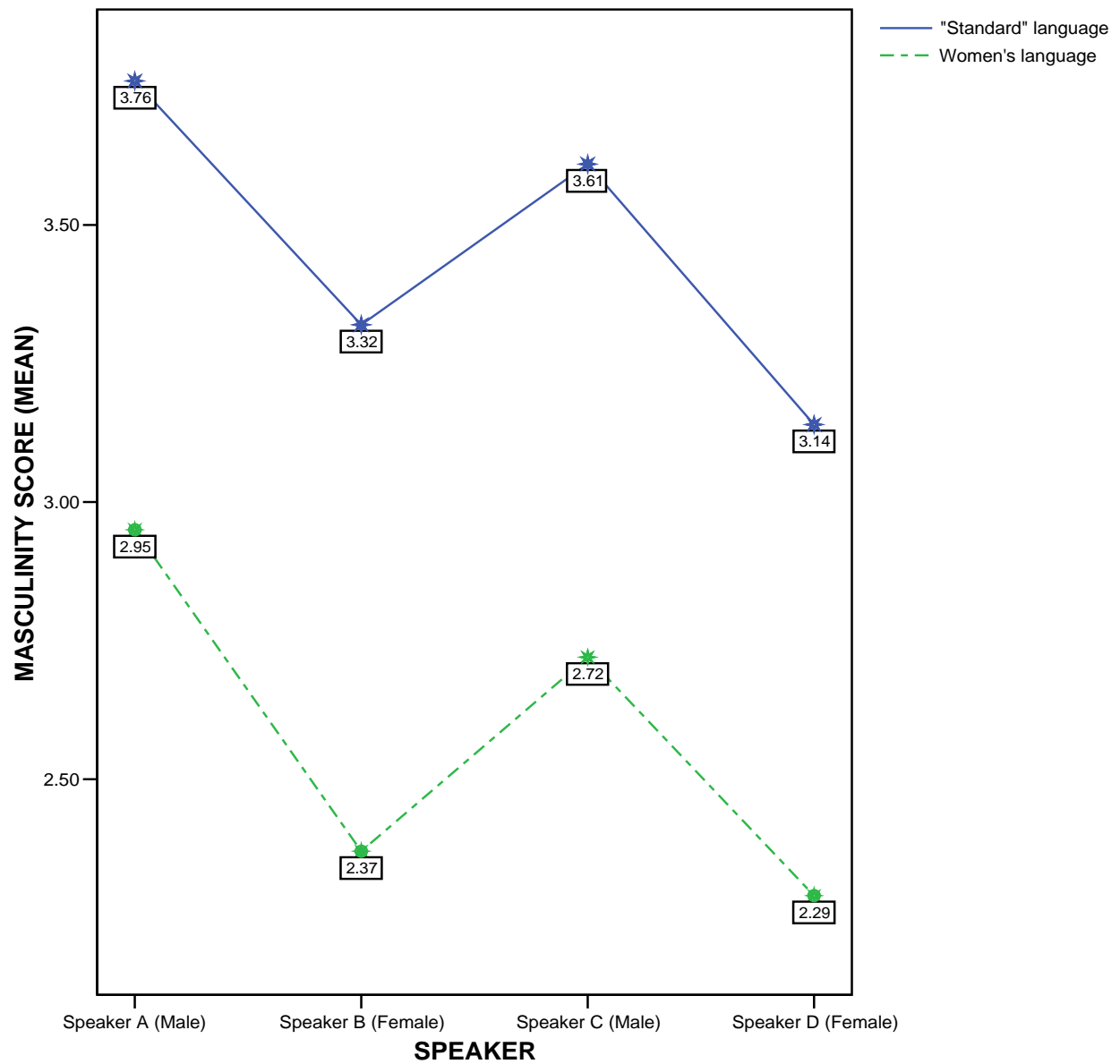


Figure 1: Comparison of Masculinity Scores for Each Speaker by Language Style

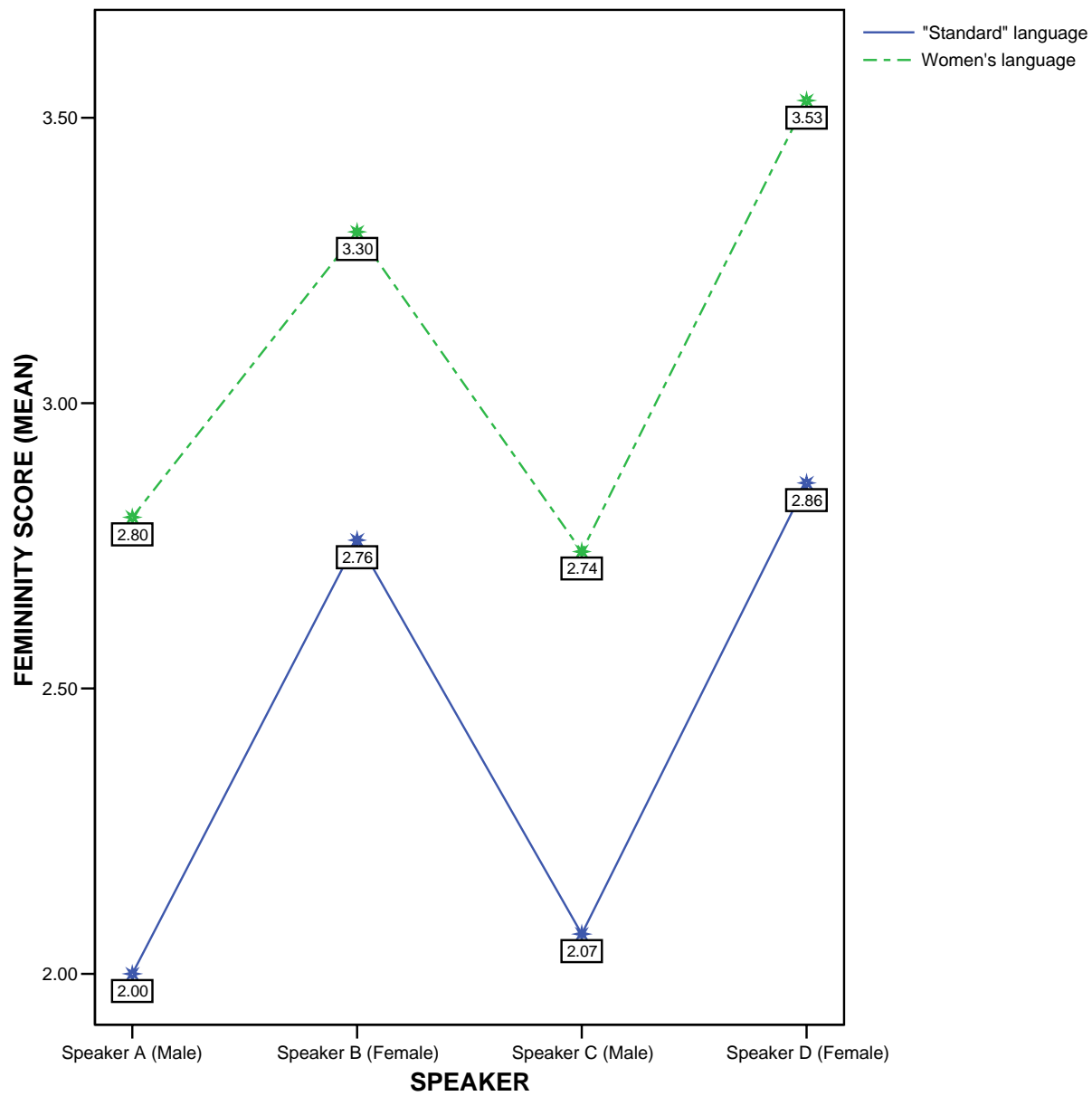


Figure 2: Comparison of Femininity Scores for Each Speaker by Language Style

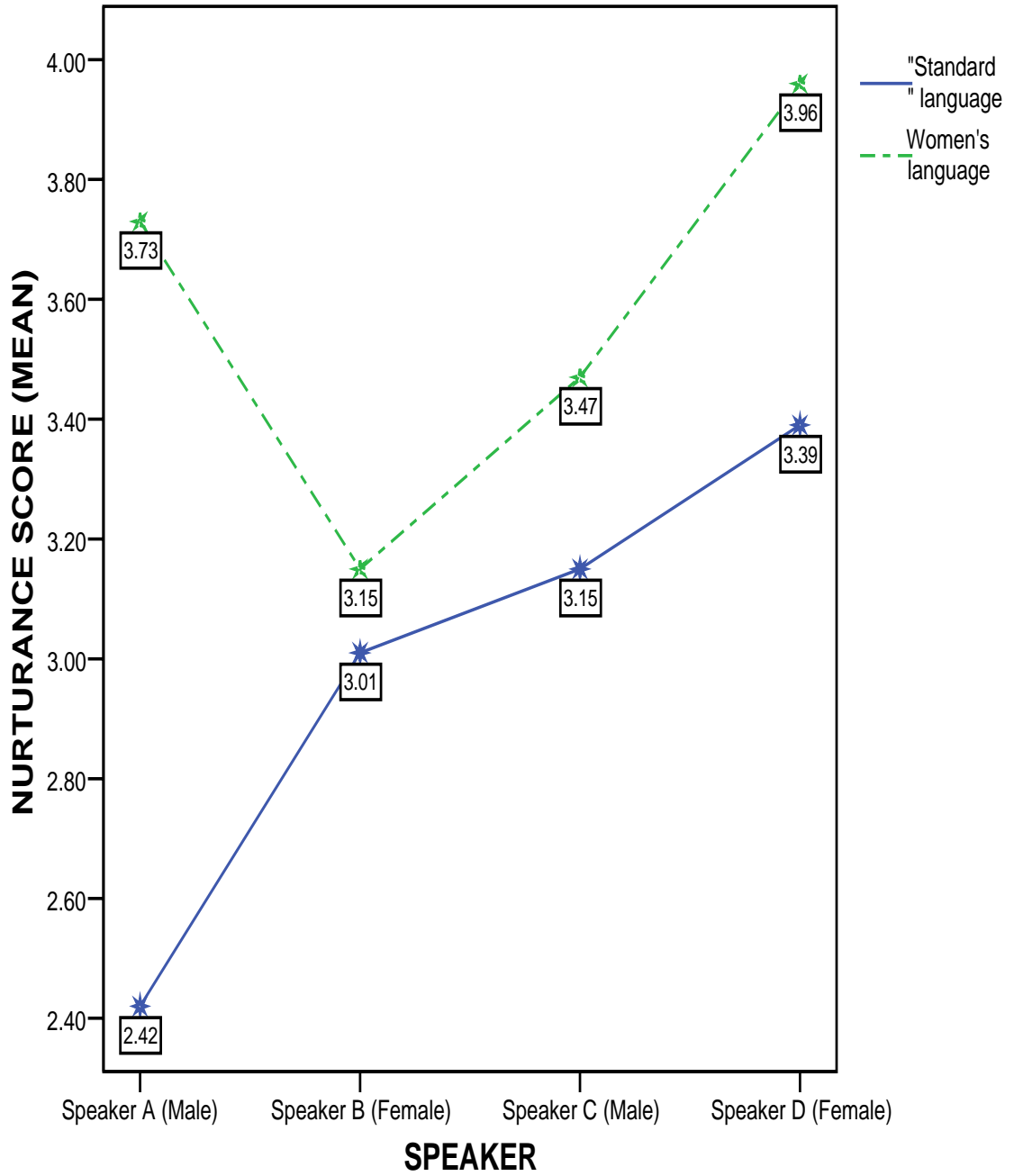


Figure 3: Comparison of Nurturance Scores for Speakers Based on Language Style

In addition to a comparison of the same speaker using different language styles, t-tests were also conducted to compare different speakers using the same language style. Significant differences occurred between Speaker C and Speaker A (both male). Speaker C was perceived

as more nurturing ($M = 3.15$) than Speaker A ($M = 2.42$) when using standard language; however Speaker A was perceived as more masculine ($M = 3.00$) than Speaker C ($M = 2.72$) when using women's language. Significant differences also occurred between Speaker B and Speaker D (both female). Speaker B was perceived as more masculine ($M = 3.32$) than Speaker D ($M = 3.14$) when using standard language, but Speaker D was perceived as more nurturing ($M = 3.40$) than Speaker B ($M = 3.01$). Finally, when using women's language, Speaker D is perceived as both more feminine ($M = 3.53$) and more nurturing ($M = 3.96$) than Speaker B ($M = 3.30, 3.15$, respectively). Tables 4 and 5 summarize the results of the t-tests for different speakers using the same language style.

Table 4: Comparing Perceptions of Different Speakers (A and C) Using Same Language Style

		SPEAKER					
		A (Male)		C (Male)		t	
<u>Style</u>	<u>Dimension</u>	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>		
"Standard"	Masc.	3.76	0.66	3.61	0.63	1.77	
	Fem.	2.00	0.71	2.07	0.60	0.94	
	Nurt.	2.42	0.59	3.15	0.53	9.26*	
Women's	Masc.	2.95	0.59	2.72	0.48	3.214	
	Fem.	2.80	0.64	2.80	0.65	0.86	
	Nurt.	3.73	0.49	3.47	0.53	4.33*	

* $p < .001$

Table 5: Comparing Perceptions of Different Speakers (B and D) Using Same Language Style

<u>Style</u>	<u>Dimension</u>	SPEAKER					
		B (Female)		D (Female)			
		<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	<u>t</u>	
“Standard”	Masc.	3.32	0.63	3.14	0.57	2.23	
	Fem.	2.77	0.55	2.86	0.57	-1.34	
	Nurt.	3.01	0.66	3.39	0.62	-4.51*	
Women’s	Masc.	2.37	0.70	2.29	0.51	0.93	
	Fem.	3.30	0.66	3.53	0.52	-3.02	
	Nurt.	3.15	0.60	3.97	0.65	-9.12*	

* $p < .001$

Again, the matched guise, for the purposes of this study, was successful as shown by the significant differences found between recordings of the same speaker using different language styles. Ideally, there would also be no differences between two speakers of the same gender using the same language style and yet the results show that this is not the case. This is not surprising, though, as the content of each of the recordings was completely different, including the features of women’s language that were used by the speakers. Thus, the comparisons made were not entirely valid as the items being measured against each other were not exactly comparable in the first place. Essentially, the recordings of women’s language contain similar counts of features considered to be components of that style, and the recordings of standard language are similar in their lack of such features (refer to Table 1). It was determined in early stages of the study that it was more important to create authentic sounding recordings for which

the speakers created their own message than to completely control for every factor in each recording.

A number of factors could have led to the differences in the results between male speakers and female speakers using the same language style. Content may have played a role. In an informal discussion with a group of participants after the data were collected, it was mentioned that a male speaker was viewed as “gentler” and less masculine simply because he mentioned having a cat⁶. Additionally, the features of women’s language may have different meanings or functions for listeners, so it may be useful in future research to only focus on one or two features of women’s language to see how they are perceived. Similarly, other features not addressed in this study could possibly have influenced an evaluation, such as voice quality (pitch, breathiness) and pronunciation style (/s/ vs. /□/). To reiterate, because authenticity of the stimuli was most important for the goals of this study, it was not vital for speakers of the same gender using the same language style to be evaluated similarly. What is most important here is that each speaker was rated differently depending on the language style used, which suggests that the language does have an effect on how a person is perceived. It would certainly be interesting if, in a future study, as many factors as possible could be controlled for to determine if speakers of the same gender receive similar evaluations. However, it may be a futile task as the factors that play a role in how a person perceives others cannot ever be entirely anticipated or controlled. And it is important to emphasize here that this study is not claiming to know everything that was involved in the evaluations of these speakers. Rather, the objective is simply to shed light on the extent to which gender stereotypes could play a role.

⁶ In actuality, none of the speakers mention a cat, but one of the men mentions a “really small dog”, which is probably what the participant is referring to. However, this is an interesting observation because the participant’s evaluation may very well have been triggered by the use of the intensifier “really”, a feature that is considered to be characteristic of women’s language.

4.4 DIMENSIONS SCORES

4.4.1 Masculinity

Analyses of masculinity scores yielded a significant main effect of speaker gender $F = 86.39$, $p < .001$ and language style, $F = 353.56$, $p < .001$. A 2x2 ANOVA revealed no significant interaction of speaker gender and language style. Results indicating differences in perception based on an interaction between speaker gender and language style were not reliable ($p > .10$). A summary of the ANOVA results is shown in Table 6.

Table 6: Summary of Two-way (speaker gender X speaker language style) ANOVA on Masculinity Scores

	<u>SS</u>	<u>df</u>	<u>MS</u>	<u>F</u>
Speaker gender	20.53	1	20.53	86.39*
Error	20.67	87	0.24	
Speaker language style	67.59	1	67.59	353.56
Error	16.63	87	0.19	
Gender X Lang. Style	0.06	1	0.06	0.34
Error	16.66	87	0.19	

* $p < .001$

As hypothesized, male speakers were perceived as more masculine ($M = 3.26$) than female speakers ($M = 2.78$) overall (that is, regardless of language style used). However, contrary to the hypothesis, the results do, in fact, reveal an effect of language style. Speakers using standard language were perceived as more masculine ($M = 3.46$) than speakers of women's

language ($M = 2.58$), regardless of speaker gender. All mean scores and standard deviations for the masculinity factor are shown in Table 7. Figure 4 illustrates the comparisons for the masculinity scores.

Table 7: Means and Standard Deviations of Masculinity Scores by Speaker Gender and Language Style

	“Standard”		Women’s		“Std.” & Women’s
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	
Male (M)	3.69	0.50	2.84	0.41	3.26
Female (F)	3.23	0.47	2.33	0.45	2.78
M & F Combined	3.46		2.58		

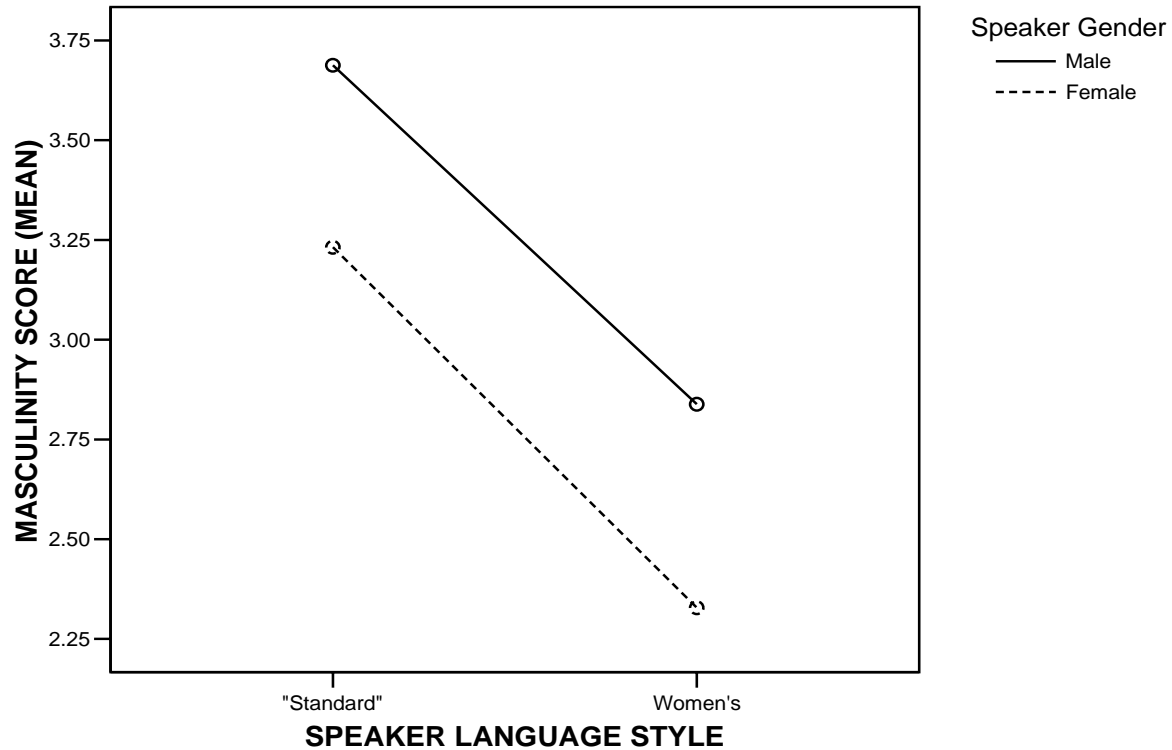


Figure 4: Comparison of Masculinity Scores by Gender and Language Style

4.4.2 Femininity

Analyses of femininity scores revealed a significant main effect of speaker gender $F = 178.82$, $p < .001$ and language style, $F = 244.14$, $p < .001$. A 2x2 ANOVA showed no significance in terms of interaction between speaker gender and language style ($p > .50$). Table 8 shows a summary of the ANOVA results.

Table 8: Summary of Two-Way (Speaker Gender X Speaker Language Style) ANOVA on Femininity Scores

Source	<u>SS</u>	<u>df</u>	<u>MS</u>	<u>F</u>
Speaker gender	44.39	1	44.39	178.82*
Error	21.60	87	0.25	
Speaker language style	39.56	1	49.56	244.14*
Error	14.10	87	0.16	
Gender X Lang. Style	0.39	1	0.39	2.45
Error	13.74	87	0.16	

* $p < .001$

In accordance with the hypothesis regarding the effect of gender, female speakers were perceived as more feminine ($M = 3.11$) than male speakers ($M = 2.40$) overall (that is, regardless of language style used). Just as the results of analysis on the masculinity dimension show, the hypothesis that language style has no significant effect on perception was not confirmed in the results. Speakers of women's language were actually perceived as more feminine ($M = 3.09$) than speakers of standard language ($M = 2.42$), regardless of speaker gender. Table 9 shows all

mean scores and standard deviations for the femininity factor. Figure 5 illustrates the comparisons for the femininity scores.

Table 9: Means and Standard Deviations of Femininity Scores by Speaker Gender and Language Style

	“Standard”		Women’s		“Std.” & Women’s
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	
Male (M)	2.03	0.53	2.77	0.54	2.40
Female (F)	2.81	0.46	3.41	0.46	3.11
M & F Combined	2.42		3.09		

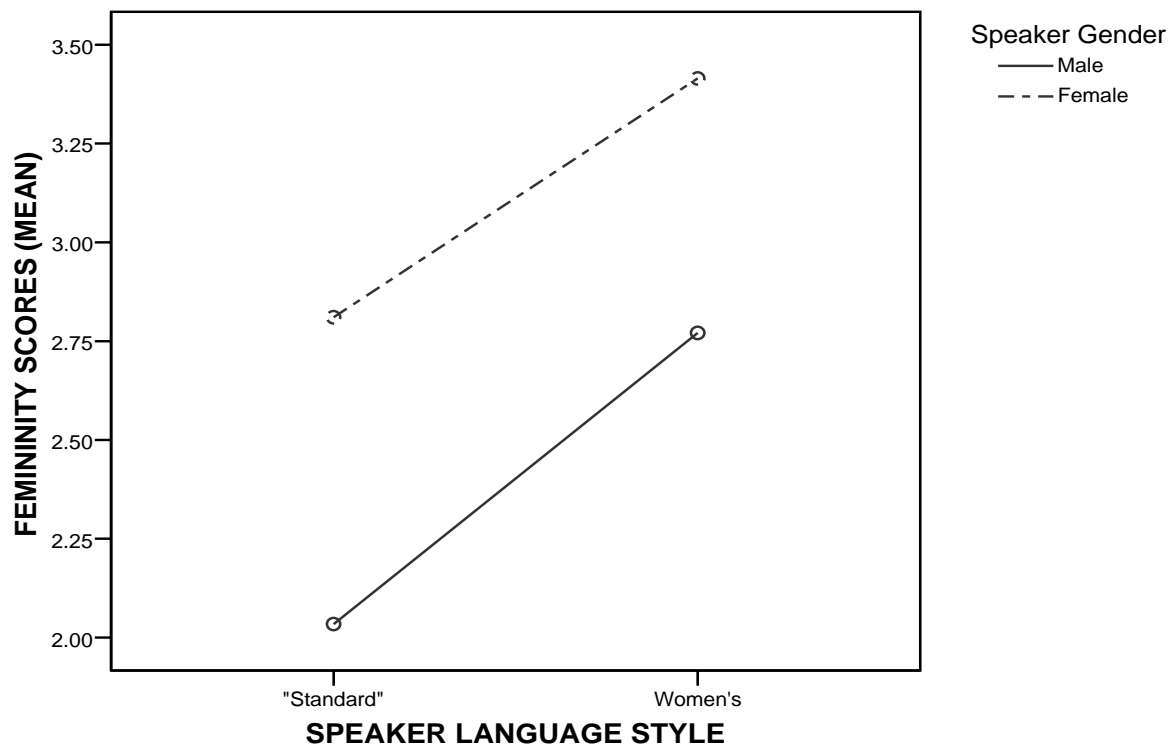


Figure 5: Comparison of Femininity Scores by Gender and Language Style

4.4.3 Nurturance

Analyses of nurturance yielded a significant main effect of speaker gender $F = 16.64, p < .001$ and language style, $F = 141.088, p < .001$, as well as a significant interaction of speaker gender and language style, $F = 42.93, p < .001$. Differences in perception based on speaker gender did, in fact, vary according to language style used. A summary of the ANOVA results is shown in Table 10.

Table 10: Summary of Two-way (speaker gender X speaker language style) ANOVA on Nurturance Scores

<u>Source</u>	<u>SS</u>	<u>df</u>	<u>MS</u>	<u>F</u>
Speaker gender	3.13	1	3.13	16.64*
Error	16.34	87	0.19	
Speaker language style	30.24	1	30.24	*141.09
Error	18.65	87	0.21	
Gender X Lang. Style	4.67	1	4.67	*42.96
Error	9.44	87	0.11	

* $p < .001$

Overall, women – regardless of language style – were evaluated as being more “nurturing” ($M = 3.38$) than men ($M = 3.19$). Speakers of women’s language were also ranked more highly ($M = 3.58$) in terms of the nurturance dimension than those who used standard language ($M = 2.99$). The interaction of speaker gender and language style reveals that female speakers were perceived as more “nurturing” ($M = 3.20$) than men ($M = 2.79$) when both used standard language; however, male and female speakers were perceived quite similarly in terms of

nurturance when both used women's language ($M = 3.60$ and 3.56 for men and women, respectively). Mean scores and standard deviations for the nurturance dimension are shown in Table 11.

Table 11: Means and Standard Deviations of Nurturance Scores by Speaker Gender and Language Style

	“Standard”		Women's		“Std.” & Women's
	<u>M</u>	<u>SD</u>	<u>M</u>	<u>SD</u>	
Male (M)	2.78	0.42	3.60	0.42	3.19
Female (F)	3.20	0.50	3.56	.50	3.38
M & F Combined	3.00		3.58		

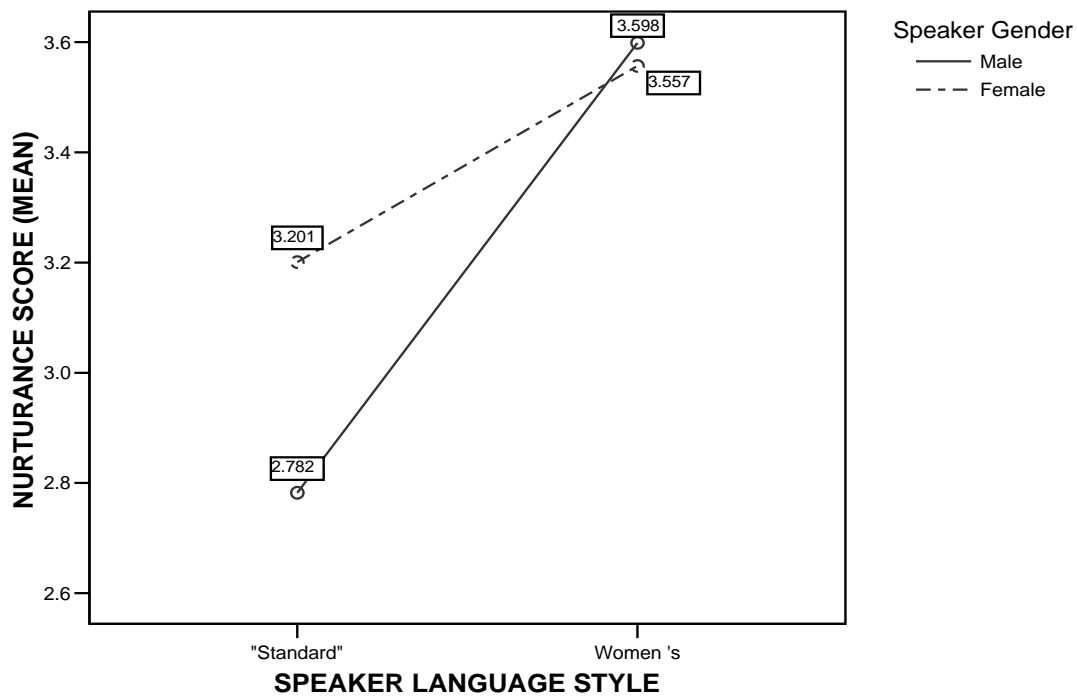


Figure 6: Comparison of Nurturance Scores for Speaker Gender by Language Style

The research questions posed at the onset of this study are as follows:

Research Question #1: Does a listener's evaluation of a speaker demonstrate an awareness of stereotypically gendered language characteristics (such as women's language)?

Research Question #2: Would a woman using women's language be subject to the same evaluation as a man using women's language?

In terms of Question #1, if a listener were aware (on some level) that the use of women's language contributed to the positioning of the speaker as uncertain or weak, it would be reasonable to imagine that this awareness would affect how a listener perceives any person, man or woman, using that type of language. In fact, the results of this study seem to support that idea. For each of the three dimensions, the evaluations indicate awareness on the part of the speakers of the gendered function of language. The fact that there is a difference in scores between women's language and standard language for the same speaker suggests that listeners consider the manner in which a person speaks (instead of only focusing on gender) when they evaluate his/her personality. It is quite possible that the participants were considering (again, on a subconscious level) what it means when a person (man or woman) speaks using certain linguistic forms (tag questions or hedges, for example).

On the other hand, the results also show a strong effect for speaker gender on listeners' evaluations. Women overall were rated as less masculine, more feminine, and more nurturing than men, regardless of the linguistic style they use. These results confirm the first hypothesis presented in this paper regarding the effect of speaker gender. It was expected that speaker gender would have the most significant effect on how the speakers are perceived, and, in fact, results appear to show that women were evaluated a certain way simply because they were women. The findings fall in line with Strand's theory regarding the categorization of speech input. Just as that study demonstrated the strength of the effect of gender on how a person is

perceived (information that did not fit gender expectations was ignored), the results of this research show that hearing a woman's voice may be enough for a person to create an image of how that person behaves. This suggests that when a participant heard a woman speaking, that input was quickly "catalogued" (as Strand argues) using gender as the first cue. The listener then applied to his/her evaluation all of the socialized expectations regarding how a woman should speak or the qualities a woman should possess.

Studies like Strand's (1999) demonstrate how focusing solely on a speaker's gender can skew a listener's perception; however, while the results of this study show a strong effect for gender, they also provide evidence that more than just gender is being taken into account when a speaker's personality is evaluated. In terms of masculinity, both male and female speakers of standard language were rated highly. Similarly, both male and female speakers of women's language were rated highly in terms of femininity. Nurturance scores show that, when using women's language, both men and women are rated highly. If language style had no bearing on how a speaker was evaluated and gender was the only influencing factor, then we would expect to see, for example, results in which men are rated similarly when using each language style; however, we see that men using women's language are not considered to be as masculine as men using standard language.

Such results offer the possibility that linguistic features can have similar meaning for both men and women as speakers (meaning they use them for the same reason), contrary to what has been argued in studies like Fishman (1980) which claimed that women and men used the same feature (e.g., minimal responses) for very different reasons. For the listeners, hearing a person use women's language (or, at least, using the features included in this experiment) led to both a more feminine, more nurturing and less masculine evaluation of the speaker's personality.

This idea coincides with the proposal by O'Barr and Atkins (1980) to refer to women's language as "powerless language" because of the common purpose it seems to serve for the men and women that use it. Erickson *et al.* (1978) also focuses on women's language/powerless language and incorporates perceptions. Erickson's results show that, just as in this study, speakers are evaluated similarly when using this style, regardless of their gender. Thus, it is also necessary to consider the context and purpose of the speech act when thinking about the function or meaning a linguistic feature may hold. A person using women's language on the witness stand in a courtroom may be perceived as powerless, whereas a person using women's language on an answering machine message inquiry about an apartment may be perceived as polite. Or, in terms of the recordings used for this study, the speakers could be considered to be in a powerless position – they have a need and are hoping the landlord they are calling will fulfill that need. In order to get what they need, the speaker must put him/herself at the mercy of the landlord, who can decide whether or not to return the call and, ultimately, allow the speaker to rent the apartment. Again, this is up for interpretation, but it is obvious that there are a number of possibilities. Generally speaking, though, the consensus among these studies seems to be that both men and women can be subject to similar evaluations based on their language style. Men and women do not necessarily use language differently, and they are not always perceived differently simply because of the types of linguistic features they use in their speech.

In addition to awareness of the variety of linguistic functions of women's language, it is also worth questioning whether or not the very concept of women's language even "holds up" (or whether it ever did). Lakoff made her argument for women's language in 1975, during the height of the women's movement and a time when it could be argued that gender roles were much more clearly defined. Women were fighting to break free from the limiting duties assigned

to them (as wives in the kitchen, as secretaries, as mothers) and to claim a place next to men as decision-makers, bosses, and money-makers. Thus, it would not be unusual for men and women to be perceived differently based on how they speak and for men and women to be encouraged to speak in different manners. After all, they were considered to be different creatures, an idea which, up until then, had rarely been questioned. Children were socialized from birth to know “what boys do” and “what girls do”.

While gendered socialization is certainly still a large part of our culture, it seems that the boundaries separating men’s and women’s roles have blurred. Because of historical movements like feminism, men and women are not expected to behave in quite the same ways that they did decades ago. A perfect example is the rise in the number of “stay-at-home dads” in the United States. In its first-ever report on stay-at-home parents, the U.S. Census Bureau reported that, in 2003, 98,000 men stayed at home to raise children while their wives worked outside of the home, a much larger percentage than in the previous decade. The report revealed an 18% increase since 1994 in the number of children who were cared for by their father while their mother went to work. The fact that this phenomenon was not even addressed in census assessments prior to 2003 implies that stay-at-home fatherhood was an even rarer occurrence thirty and forty years ago. It also demonstrates that the blurring of the gender role boundaries cannot be ignored. This leads to the question that if it is more acceptable for men and women to behave similarly, why couldn’t this also apply to the use of language? It would certainly explain the patterns in the data. One could argue that the listeners, who were mostly of a younger generation (approximately 21 years of age), were more willing to perceive men in terms of their femininity, and to rating men and women similarly regarding masculinity and/or femininity. By having

more exposure to the blurring of gender roles, people may be less inclined to attach a specific meaning to a language style simply because the person using it is a man or a woman.

In a similar vein, Research Question #2 asks whether a man using women's language would be evaluated similarly to a woman using women's language. Neither the masculinity nor femininity scores demonstrate significant differences in how men using women's language and women using women's language are perceived. There is also a lack of significance in the differences between men using standard language and women using standard language. The only dimension for which men and women were rated similarly was nurturance when both genders used women's language. Women, in general, were considered to be more nurturing than men. Interestingly, though, the difference is very small between men and women both using women's language. Men and women were rated similarly in terms of these nurturing traits when they both used women's language. Why could this be? A possible explanation for this interaction can be found in the factor analysis that extracted the nurturance dimension. As the results of the analysis show, the adjectives that make up this dimension (sensitive, warm, understanding) did not fit into the masculinity-femininity framework created by the statistical analysis. It could be postulated that these adjectives have less of a gendered meaning for listeners, even though, according to the BSRI, these adjectives are indicators of femininity. Perhaps women's language, instead of positioning a speaker as more feminine, or tentative and uncertain, positioned both male and female speakers in such a way that listeners considered them to be similarly warm, understanding, and sensitive (nurturing). Again, this suggests that the features of women's language do, in fact, have many functions.

Another possibility is that the nurturance dimension is actually measuring another type of femininity. Similar to what the various factor analyses revealed, as discussed in Brems and

Johnson (1990), it may make more sense to regard the concept of femininity as multi-dimensional. For example, it could be more useful to consider the adjectives used in the femininity dimension – *childlike*, *shy*, and *feminine* – as the more socially undesirable manifestations of femininity. On the other hand, the “nurturing” adjectives – *understanding*, *warm*, and *sensitive* – could be viewed as more desirable qualities of femininity and, as a result, they are more acceptable for a man to possess. This would explain the similarity in scores in the nurturance dimension between women using women’s language and men using women’s language. This style of language in general may make the speaker sound warmer, more understanding, and more sensitive, and listeners may have been more inclined to rate a man similarly to a woman in terms of these adjectives.

4.5 EFFECT OF PARTICIPANT GENDER

The effect of the participants’ gender was also examined to see if men and women rated speakers differently. Various t-tests were conducted to determine if the gender of the participant had an impact on the perception of speakers. In general, there were no significant differences between men’s and women’s evaluations of the four conditions (men using standard language, men using women’s language, women using standard language, women using women’s language) in terms of the three dimensions. These results bear out the hypothesis that men and women would perceive speakers with no significant difference. However, all of the results are not as straightforward. Within the masculinity dimension, two groups had significant ($p < .01$) results: men using women’s language as perceived by women ($p = .003$) when analyzed with equal

variances assumed, and women speaking standard language as perceived by men ($p = .042$) when analyzed with equal variances not assumed. Figure 7 shows these results.

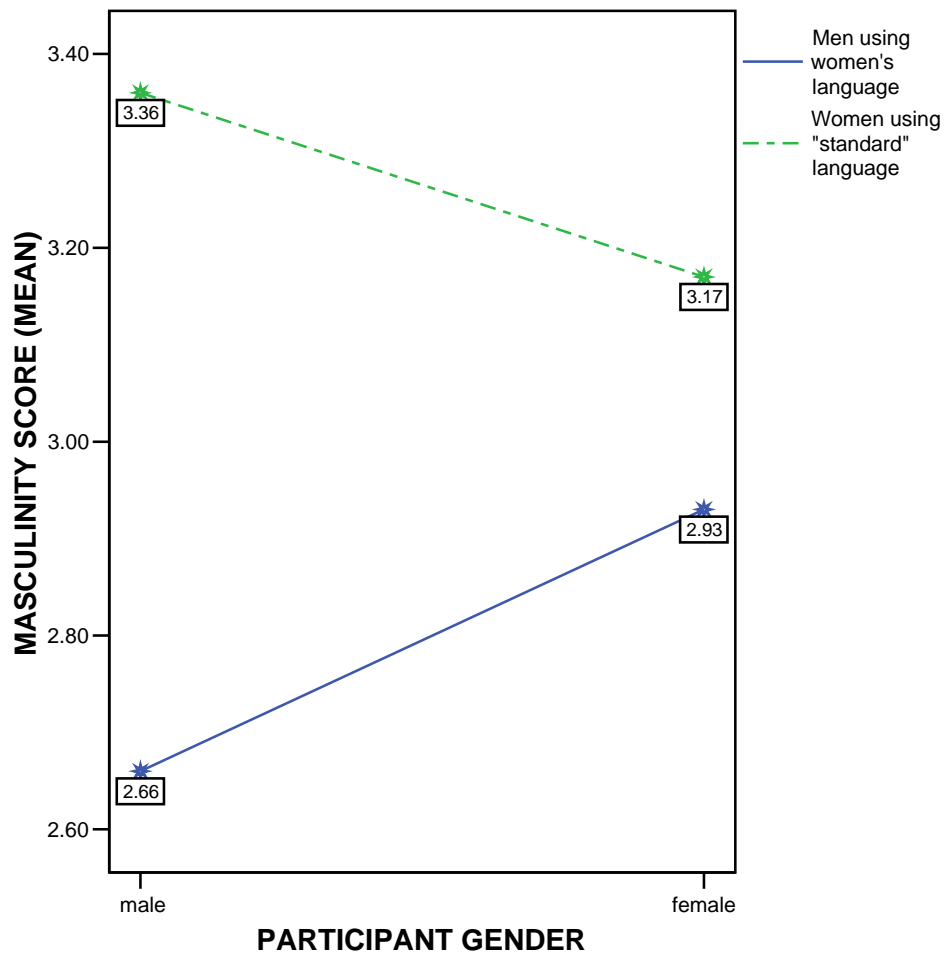


Figure 7: Comparison of Masculinity Scores for Men Using Women's Language and Women Using "Standard" Language Based on Participant Gender

As the chart illustrates, men using women's language were perceived by women to be more masculine, and women using standard language were perceived by men to be more masculine. These results are remarkable because for both groups of speakers, the language style used is what would be considered to be atypical. Women are "supposed" to use women's language, men are "supposed" to use a more standard style, but in these cases they are both using

the style that is least often associated with their gender. Furthermore, they are being perceived as more masculine when using this style. This suggests that women were more apt to evaluate a man as masculine, despite using women's language. So, for women, the gender of the speaker seems to have more of an effect than language style on their perception. If the argument were made that women were more frequent users of women's language, then these results would also suggest that perhaps women are less aware of the alleged gender-specific function these linguistic features have. Another possibility is that it is more socially acceptable for women to acknowledge feminine traits in men, whereas for men, the stigma attached to being perceived as "feminine" would lead men to disregard or ignore feminine behavior displayed by men. Quite possibly, a man using a more feminine style of speech would have no effect on how a woman interprets his masculinity.

For men, women using standard language were perceived as more masculine, which implies that, contrary to what the women's evaluations revealed, male listeners might have more specific ideas regarding how women are supposed to speak. Hearing a woman use a language style that does not fit with their expectations could have led the men to rate the female speakers as more masculine. Similar to the hypothesis regarding social stigma, it could also be that men are less willing or able to consider gender outside of their rigid stereotypical expectations regarding behavior. Thus, they do not acknowledge when men behave in a feminine manner and when women behave outside of the expected feminine manner, they are perceived as more masculine (this would assume that masculinity is undesirable for women). Of course, all of these interpretations are mere speculation, and somewhat risky speculation at that, given the assumptions being made about social stigma and "typical" behavior. However, the results –

which were unexpected - bring up interesting questions, and further investigation would be needed to test these ideas out.

5.0 CONCLUSION

The results of this study have significant implications for language and gender research. The inclusion of listener attitudes provides information that much of the past research has not emphasized, namely, how gender and language interact to create an attitude toward a speaker. This suggests that the construction of a social identity is actually a cooperative process, incorporating the language style a speaker chooses (and whatever meaning these choices have for the speaker) as well as the listener's interpretation of that language style (in addition to other non-linguistic factors). All of these elements work together as we initiate, maintain, or terminate social relationships. The observations made in this study confirm that associating a specific language style with a certain type of person (a man or woman, in this case) is indeed an inadequate approach. While the results revealed the undeniable and overarching influence of gender on how a person is evaluated (after all, men were rated as more masculine and women as more feminine no matter how they spoke), they also show that listeners are aware (on some level) of the different functions that language can have. The use of women's language does seem to be associated with a more feminine, less masculine, and more nurturing personality (regardless of gender), as the evaluations demonstrated. And, although it may seem that these two conclusions are at odds, in actuality they are not. Rather, by showing that both gender and language style affect attitudes, they underscore the importance of considering a "bigger picture" when we interpret language use. These "conflicting" results expose what was lacking in the

research methods utilized in past. Lakoff's and Fishman's arguments, for example, only tell part of the story and do not allow for multi-faceted interpretations.

In broader terms, the goal of this study was to demonstrate the importance of incorporating many types of interpretations when considering the functions of language. Useful factors to incorporate in analysis include: linguistic form or feature used, context (local and global), participant background, and, of course, listener attitudes. This study did not incorporate all of these factors; however, it did demonstrate that the link between language and gender is not "black and white" and that it would be unwise not to acknowledge the influence that other have when we interpret a person's speech style. Still, the results of this study reveal that speaker gender has a significant effect on a listener's attitude toward that speaker, which reminds us that gender is arguably one of the most salient and powerful components of one's social identity.

And while the studies of the past are useful and have paved the way for current research, many are simply not expansive enough. To be able to gain a richer understanding of language variation and the functions of linguistic forms, various levels and dimensions of interpretation must be considered.

5.1 DIRECTIONS FOR FUTURE RESEARCH

Perhaps the most challenging issue for a study like this is controlling for the various factors that can influence evaluation, which is an important issue to address in future research. As discussed in the previous section, it is virtually impossible to control for all of the factors that have the potential to play a role in how a person perceives a speaker. However, it is also crucial to consider whether or not it is actually necessary to control for such issues. The results of this

study may have been different had factors such as recording content (in terms of what the speakers said and the linguistic features used to say it) been more explicitly defined. Whereas this study emphasized a more natural and authentic approach with the goal of eliciting genuine and somewhat impulsive evaluations from participants, it will be interesting to observe the results of a task in which all of the speakers use the same type and amount of linguistic features (for example, no more and no less than three tag questions and six hedges per women's language recording) and followed a pre-constructed script. On the other hand, addressing other factors not controlled in this study – speaker voice quality and variations in pronunciation, for example - would probably not have much of an effect on the overall results, as the interpretation of these things can vary so greatly from person to person. Still, future studies would benefit by incorporating a more streamlined matched guise technique that controls for more factors (like those mentioned at the beginning of this section). Doing so would result in a more precise understanding of the relationship, if any, between gender, language style, and evaluations of speakers.

There are also some practical issues to consider. Based on informal discussions with participants, it was revealed that the wording of the ratings for the questionnaire may have been problematic. Instead of determining how *often* the speaker demonstrates the various personality traits (*never*, *rarely*, *sometimes*, etc.), it may make more sense to ask listeners to rate how well the adjective describes the person (for example, *never true* or *almost never true*). This wording does not require the listener to think about how the speaker behaves generally; rather, the listener can just consider how they would describe that person at that particular moment (while they are hearing the speaker talk).

Finally, because this study focuses on listener attitudes, it would be beneficial to include a qualitative element so as to glean as much information as possible from participants regarding their evaluations. Originally, the survey used in this study was to include a section for open-ended comments from the participant regarding anything that came to mind during the survey. Unfortunately, due to time constraints and the amount of data collected (which was more than expected), this section was omitted from the survey. However, the informal discussions with participants (which occurred in two sections of Introduction to Linguistics at the request of the professor) revealed interesting ideas and observations from participants. If a similar study were done in the future it would be useful to include an opportunity to write personal comments about the survey or to participate in a discussion about the task.

APPENDIX A

RECORDING TRANSCRIPTIONS

Transcription Conventions⁷

<u>Symbol</u>	<u>Meaning</u>
:	elongated vowel sound
↑	rising intonation
↓	falling intonation
→	speaker's turn continues without interruption
..	pause shorter than one second
...	pause longer than one second
((xxx))	indicates manner in which utterance is spoken [for example, ((quietly))]
ALL CAPS	utterance is stressed/emphasized

Recording #1: Standard language (Speaker C, male)

- 1) Yes hello my name is Lou Adams and I'm calling concerning your one bedroom →
- 2) apartment in Lawrenceville that's listed on the first ↑ floor →
- 3) If you could give me a call back I would appreciate it I have some questions →
- 4) concerning the utilities if uh-any of those are covered in the rent →
- 5) and also how close the um..apartment itself is to the ↑ bus↓ lines I will be available →
- 6) between three and four o'clock this after ↑ noon and again if you could give me a →
- 7) ↑ call I would ap↑pre↓ciate it

⁷ Transcription conventions were adapted from Tannen (1989a) and Jefferson (1979)

Recording #2: Standard language (Speaker B, female)

- 1) Hi my name is Tina...I'm calling about the: North Oakland one bedroom apartment listed →
- 2) in this week's City Paper →
- 3) Before I decide to a look at it though I need to know what floor the apartment is on →
- 4) AND what your policy is on pets →
- 5) Please call me back as soon as possible ↑Thanks

Recording #3: Women's language (Speaker D, female)

- 1) Hi:: my name is Mela↑nie and I'm calling about the apart↑ment in Oak↑land →
- 2) I saw your ad in the pa↑per and I'm very interes↑ted I'd like to find out more about ↑ it →
- 3) like if there's any way I could have a ↑ tour that would be SO helpful →
- 4) I'm really interested in the one bed↑room If someone could get back to me in the next few →
- 5) ↑days I'd REALLY appreciate it ↑↓Thanks

Recording #4: Women's language (Speaker B, female)

- 1) Hi..my name's Jess and I'm calling about the one bedroom apartment in Shady↑side you →
- 2) have listed in the City Paper →
- 3) Uh-umm I don't know if it would still be open but I was ↑kinda hoping to get a little bit →
- 4) more information →
- 5) I ↑know it says Shadyside so this could be a really stupid question, but..I need to know →
- 6) whether there's a bus route nearby →
- 7) Um...also if you could call me back as soon as possible, I kinda need to find a place by →
- 8) the end of the month Thanks

Recording #5: Standard language (Speaker A, male)

- 1) Hel↓lo this is ↑↓Carl I'm calling about the one bedroom apartment in Green↓field →
- 2) Uh I have two questions the first is about whether or not a month to month lease →
- 3) might be possi↑ble the second is uh...to what degree utilities are included in →
- 4) the 450 dollar rent..charge It's best to reach me during the day →
- 5) ((quietly)) thank you.

Recording #6: Women's language (Speaker C, male)

- 1) Hello my name is ↑↓Lou and I'm calling about um your one bedroom apartments in South →
- 2) Side Flats if you haven't rented them already →
- 3) Um I'd kinda like to come by and take a LOOK at them I also had some questions about →
- 4) um..whether or not utilities are covered if any of them are covered Um I'm available →
- 5) today between three and four o'clock in the afternoon so if you could gimme a call that'd →
- 6) be ↑good ↑Thank you

Recording #7: Standard language (Speaker D, female)

- 1) Hi my name is Michelle and I'm calling about the apartment in Shadyside →

- 2) I'd like to know about the specific location and set up a tour →
- 3) and also whether utilities are included My friend and I are looking to rent a two bedroom →
- 4) I'd appreciate a call back at your earliest convenience →
- 5) Thanks a ↑↓lot

Recording #8: Women's language (Speaker A, male)

- 1) Hi this is Kevin uh I'm..calling uh about the one bedroom apartment in Bloom↑field →
- 2) which I'm very interested in uh I have two questions the first is about park↑ing →
- 3) uh I wanted to know what might be availa↑ble uh the second is about pets uh because I →
- 4) have a REALLY small ↑dog uh and I'm hoping he won't be a ↑problem →
- 5) uh I REALLY appreciate it and thanks for getting back to me

APPENDIX B

EXAMPLE OF SURVEY ID FORM

Your ID #:_____ Listening Session:_____

Please read this entire page before beginning the survey!

Description of your task:

When we listen to a voice over the radio or on an answering machine, it is quite natural to imagine what kind of person is “attached” to that voice. This is precisely what you are to do when you listen to the recordings in this survey. After each recording, consider the list of adjectives on the screen. To what extent would you describe the person you hear using these adjectives? Think about the overall personality of the person attached to the voice you hear.

It is best not to think about your answer for too long – your first impression of the speaker’s personality is the best answer! Remember, there are no right or wrong answers on this survey.

How to access and complete the survey:

Click on the “Language Attitude Survey” icon on the desktop.

The survey will begin with a few questions about your background. If you do not answer all the questions, you will not be able to continue to the next page of the survey!

Put on the headphones attached to your PC. Play each sound file first and fill out the survey after you’ve listened to the file in its entirety. You will not be able to click on any buttons until the recording has finished.

When you are finished with the survey, click on “I am finished.” When the “Continue” button appears, click on it to be taken to the next survey and sound file. Repeat the same process explained in the step above.

There are eight surveys total. After you complete the eighth one, you will be taken to a final screen. Click on “Quit” to finish.

If you have any questions or problems while you are working, raise your hand and I will come and assist you!

BIBLIOGRAPHY

- Batstone, Susan and Tuomi, Seppo K. 1981. Perceptual characteristics of female voices. *Language and Speech*, 24(2).
- Bem, Sandra L. 1993. *The Lenses of Gender*. New Haven: Yale University Press.
- Bem, Sandra L. 1974. The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology*, 42: 155-162.
- Brems, Christiane and Johnson, Mark E. 1990. Reexamination of the Bem Sex Role Inventory: The interpersonal BSRI. *Journal of Personality Assessment*, 55(3&4): 484-498.
- Cameron, Deborah; McAlinden, Fiona; and O'Leary, Kathy. 1988. Lakoff in context: the social and linguistic function of tag questions. In Coates and Cameron 1988, 74-93.
- Choi, Namok and Fuqua, Dale R. 2003. The structure of the Bem Sex Role Inventory: a summary report of 23 validation studies. *Educational and Psychological Measurement*, 63(5):872-887.
- Clopper, Cynthia G.; Conrey, Brianna; and Pisoni, David B. 2005. Effects of talker gender on dialect categorization. *Journal of Language and Social Psychology*, 24(2):182-206.
- Coates, Jennifer and Cameron, Deborah (eds.). 1988. *Women in Their Speech Communities: New Perspectives on Language and Sex*. London and New York: Longman.
- Delph-Janiurek, Tom. 1999. Sounding gender(ed): vocal performances in English university teaching spaces. *Gender, Place and Culture*, 6:137-153.
- Delph-Janiurek, Tom. 2000. 'Walking the walk and talking the talk': bodies, conversation, gender and power in higher education in England. *Social & Cultural Geography*, 1:83-100.
- Dubois, Betty Lou and Crouch, Isabel. 1975. The question of tag questions in women's speech: they don't really use more of them, do they? *Language in Society*, 4:289-294.

- Erickson, Bonnie; Lind, E. Allan; Johnson, Bruce C.; and O'Barr, William M. 1978. Speech style and impression formation in a court setting: The effects of "powerful" and "powerless" speech. *Journal of Experimental Social Psychology*, 14: 266-279.
- Fishman, Pamela. 1983. Interaction: the work women do. In Thorne, Kramarae, and Henley 1983, 89-102.
- Glass, Gene V. and Hopkins, Kenneth D. 1996. Statistical Methods in Education and Psychology, Third Edition. Boston: Allyn and Bacon.
- Harris, Allen C. 1994. Ethnicity as a determinant of sex role identity: a replication study of item selection for the Bem Sex Role Inventory. *Sex Roles*, 31(3/4): 241-257.
- Holmes, Janet. 1982. The functions of tag questions. *English Language Research Journal*, 3:40-65.
- Kramarae, Cheris. 1982. Gender: How she speaks. In Ryan and Giles 1982, 84-98.
- Lakoff, Robin. 1975. Language and Woman's Place. New York: Harper & Row.
- Lambert, Wallace. 1967. A Social Psychology of Bilingualism. In Paulston and Tucker, eds. 2003, 305-317
- McMillan, J.R.; Clifton, A.K.; McGrath, D.: and Gale, W.S. 1977. Woman's language: uncertainty or interpersonal sensitivity and emotionality? *Sex Roles*, 3(6):545-549.
- O'Barr, William M. and Atkins, Bowman K. 1980. "Women's language" or "powerless language"? In McConnell-Ginet, Borker, and Furman 1980, 93-110.
- Purnell, Thomas; Idsardi, William; and Baugh, John. Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology*, 18(1): 10-30.
- Strand, Elizabeth. 1999. Uncovering the role of gender stereotypes in speech perception. *Journal of Language and Social Psychology*, 18: 86-99.
- West, Candace and Zimmerman, Don. 1983. Small Insults: A study of interruptions in cross-sex conversations between unacquainted persons. In Thorne, Kramarae, and Henley 1983, 103-117.
- Zimmerman, Don and West, Candace. 1975. Sex roles, interruptions and silences in conversation. In Thorne and Henley 1975, 105-129.